A Role for Antisense Transcripts in Regulation of Viral Gene Expression; Gene silencing through hypermethylation of CpG islands in the HCMV genome?

A Senior Honors Thesis

Presented in partial fulfillment of the requirements for graduation with research distinction in the undergraduate colleges of The Ohio State University.

By

Mary Sivulich, B.S.

The Ohio State University 2008

Project Advisor: Dr. Joanne Trgovcich, Assistant Professor, Department of Pathology, The Ohio State University Medical Center

Work for this thesis was supported in part by the College of Arts and Sciences Undergraduate Research Scholarship and the College of Biological Sciences Dean's Undergraduate Research Fund at The Ohio State University. This project was developed in conjunction with the larger study, "Antisense Transcription in the Human Cytomegalovirus Transcriptome" conducted by Guojuan Zhang, Bindu Raghavan, Mark Kotur, Jacquelyn Cheatham, Daniel Sedmak, Charles Cook, James Waldman, and Joanne Trgovcich of the Ohio State University. Special thanks to Dr. Ilya Ioshikhes of The Ohio State University for consulting with the mapping of CpG islands for the HCMV genome, as well as to Dr. Peter Charles of the University of North Carolina and Dr. Kari Green-Church of the Ohio State University for assistance in global methylation experiments.

Abstract

Human cytomegalovirus (HCMV) infections are prevalent in human populations and may engender serious diseases, particularly in individuals with compromised or immature immune systems. The Trgovcich laboratory conducted an analysis of the HCMV transcriptome using a cDNA library approach. This analysis led to the discovery that noncoding and antisense (AS) transcripts accumulate to high levels during infection (21). Such findings support the idea that the complexity of viral gene products is greater than previously recognized. Moreover, these findings suggest that epigenetic regulatory mechanisms influence the viral life cycle. Antisense transcripts have been shown to function in regulating gene expression through transcriptional interference and masking of regulatory elements involved in splicing, localization, and translation. Antisense transcripts also serve as primary transcripts for miRNAs. In addition, they are implicated in gene silencing through methylation of CpG islands associated with gene promoter regions (11). This latter process is known to be involved in genomic imprinting and silencing of genes on the inactive X chromosome (20).

We set out to test the hypothesis that AS transcripts mediate the silencing of viral gene promoters associated with CpG islands and that this is an important mechanism in regulating the pattern of viral gene expression. CpG islands were mapped across the entire viral genome and overlapping AS transcripts were identified. It was expected that methylated regions of the genome would be resistant to cutting with methylation sensitive restriction enzymes, and that this could be observed by an alteration in the sizes of DNA fragments generated after digestion with these enzymes. The DNA fragments were measured by size separation on agarose gels combined with Southern transfer and hybridization techniques using probes specific for 5 selected genomic regions. Upon analysis, all five regions yielded evidence of patterns consistent with

unmethylated CpG dinucleotides. The major limitation of this method was the limited number of CpG dinucleotides within the genome that could be analyzed. As an alternate approach, the lab will investigate methylation of the genome using two approaches. First, we will employ a liquid chromatography and mass spectroscopic approach to directly detect and measure methylated cytosines in viral genomic DNA. If successful, we will use a modified "ChIP on chip" (Chromatin IP) approach to specifically identify methylated regions of the genome. Presumptive methylated regions will be isolated through restriction digests of viral DNA followed by immunoprecipitation with an antibody specific for 5-methylcytosine. These regions will be identified by hybridization to a tiling array of the HCMV genome. We expect these studies to provide a global view of genome methylation and definitively determine methylation status in the genome. Most importantly, by mapping the putative methylated regions, it will be possible to correlate methylated regions with the positions of the antisense transcripts. These studies will lay the groundwork for future studies to determine if antisense transcription in the HCMV genome functions to regulate CpG island methylation and viral gene transcription.

Table of Contents

Page

Tables and Figures

Page

Tables

Figures

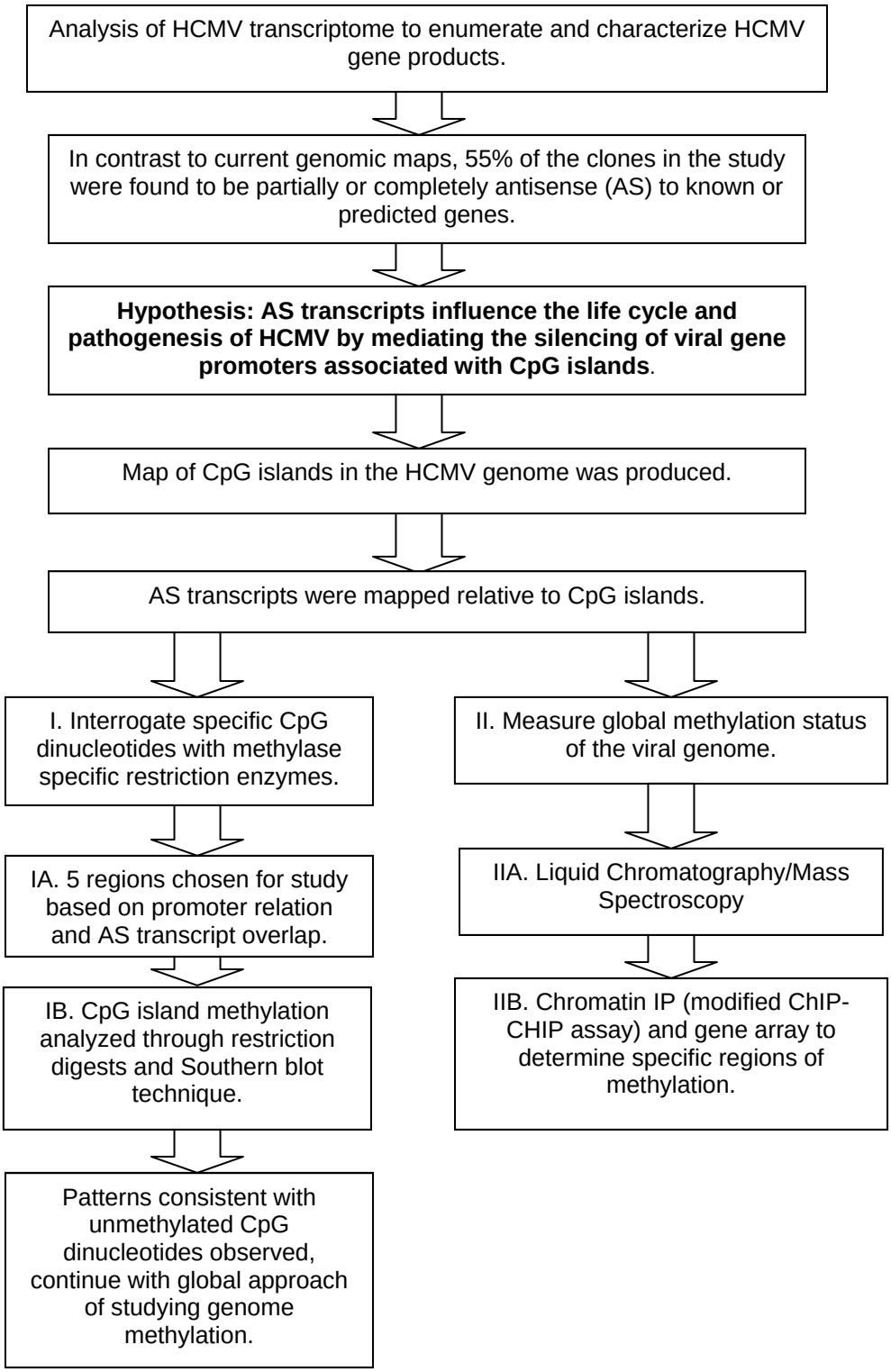
Background and Introduction

Human Cytomegalovirus. Human cytomegalovirus (HCMV) is a member of the herpesvirus family, a classification which also includes viruses that are the source of herpes (herpes simplex virus), chicken pox (varicella-zoster virus) and infectious mononucleosis (Epstein-Barr virus). It is transmitted by body fluids and once contracted establishes life-long infections in its hosts (18). These infections can persist chronically at active low levels or in quiescent latent states. The bone marrow is believed to be a major reservoir of latent virus. Most humans are exposed to the virus during their lifetime, HCMV infecting an estimated 50-85% of the population (2). Though most commonly benign, HCMV infections are often serious in those individuals with immature or compromised immune systems. In addition, the virus produces an array of proteins that disrupt innate and adaptive immune responses (4). HCMV causes congenital diseases such as growth retardations and jaundice. In immunocompetent adults, mononucleosis is the most recognized manifestation of infection. In immunosuppressed adults, HCMV is also associated with sight threatening infections such as retinitis as well as other diseases including colitis. Transplant patients with HCMV infections may suffer graft loss, vascular disease, or other manifestations including hepatitis and pneumonitis (15). In healthy individuals the virus's presence may be asymptomatic or manifest in symptoms that include fever, sore throat, fatigue, and swollen glands, indicators that make it difficult to distinguish from many other mild illnesses based on symptoms alone. Some data also suggests that HCMV may be associated with chronic disease states such as atherosclerosis, arthritis, and cancer (17).

The viral life cycle

The structure of an HCMV viral particle includes an outer lipid bilayer that contains virally encoded glycoproteins and a core containing viral DNA surrounded by an icosahedral capsid. Between the capsid and the envelope is an amorphous layer referred to as the tegument which includes proteins necessary for replication, and mRNAs (8). When the virus infects a host, the HCMV glycoproteins bind to cellular receptors, activating signaling pathways and cellular transcription factors. The virus then enters the cells releasing viral capsid, virion proteins and virion mRNA transcripts into the cytoplasm. The viral capsid is transported to the nucleus of the host cell via active transport to the nuclear pore. In the nucleus, the viral DNA is released and transcribed by the host cell machinery and viral transcription factors that are packaged in the virion. Viral gene expression occurs in a temporally regulated manner, first with expression of immediate early genes, followed by early genes, and after viral replication has been initiated, late genes. The immediate early genes are required for both early and late gene expression, while the early and late genes typically encode for replication factors and structural components of the virion respectively. The viral DNA and viral and cellular proteins and transcripts are then packaged into the virion, and an enveloped infectious viral particle is released. At the start of the next infection, the viral envelope will fuse once again with the plasma membrane of the host cell and the process will repeat, taking approximately seventy-two hours (16). The mechanisms that regulate this temporal gene expression are poorly understood, especially for late genes. Also poorly defined are the mechanisms of gene regulation during latency of the virus. **Therefore, we chose to test the hypothesis that epigenetic regulation, possibly through CpG island methylation, contributes to viral gene regulation during the lytic cycle.** The strategy we will employ to test this hypothesis is shown in Fig.1.

Figure 1 – Overview of Project



The viral Transcriptome

The HCMV genome is the largest of the herpes viruses spanning 230kb. The DNA sequence of the genome was first published in 1989, though the precise number of viral genes is still in question. In furthering their knowledge of such an expansive genome, the Trgovcich laboratory at the Ohio State University conducted an analysis of the HCMV transcriptome, obtaining sequence data on isolated clones selected from the lab's cDNA library composed of different temporal classes of the viral transcripts. As shown in Figure 2, this analysis yielded the recognition of noncoding and antisense (AS) transcripts that accumulate during infection and are not currently included on available genomic maps.

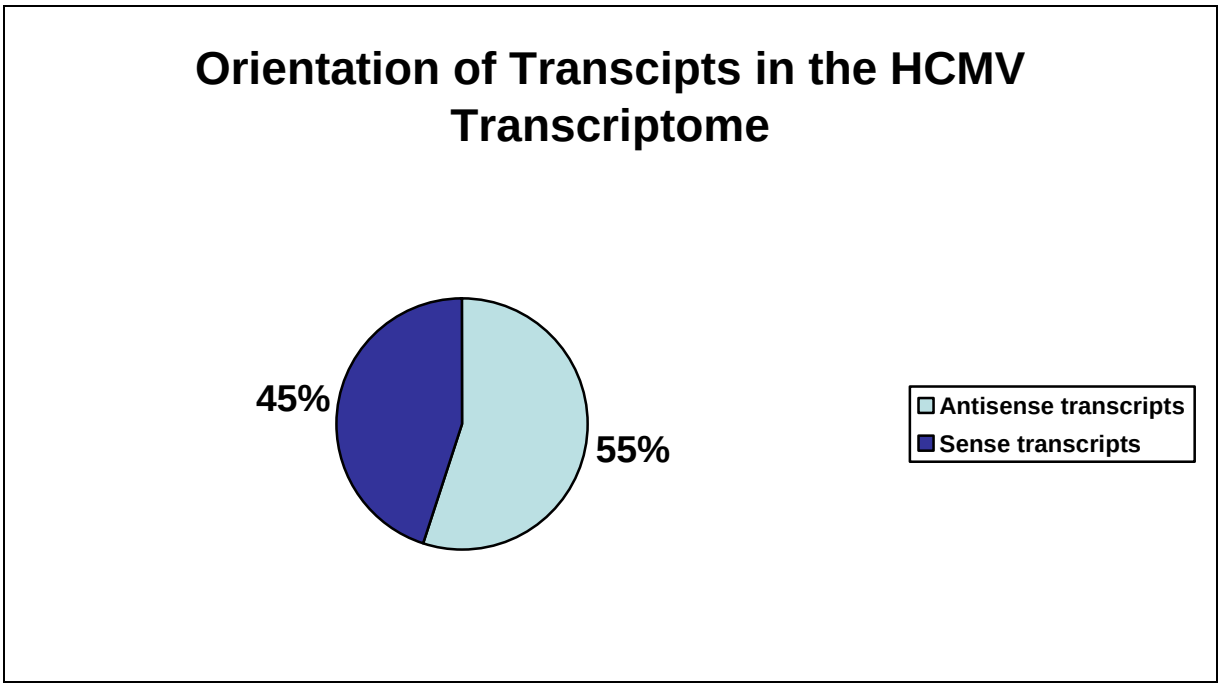


Figure 2. Prevalence of partial or complete antisense transcripts to known or predicted genes in the HCMV Transcriptome based on a recent analysis (2007) conducted by the Trgovcich laboratory.

Antisense Transcripts. Sense-Antisense (S-AS) transcripts have been found in a wide range of genomes, with a subset of S-AS pairs that are recognized as being conserved across species (21). Sense transcripts are transcribed from the strand of DNA specifying the genetic code, and are transported to the cytoplasm for translation into proteins. As Figure 3 illustrates, antisense transcripts are then synthesized from the opposite strand of DNA in the double helix. S-AS transcripts may form double-stranded intermediates through complementary base pairing (13). Antisense transcripts are classified as *cis* or *trans* in nature: *cis* transcripts are transcribed from opposite strands of the same genomic locus and are predicted to have longer and more complementary sequences for S transcripts than the *trans* AS transcripts that are derived from separate loci (11).

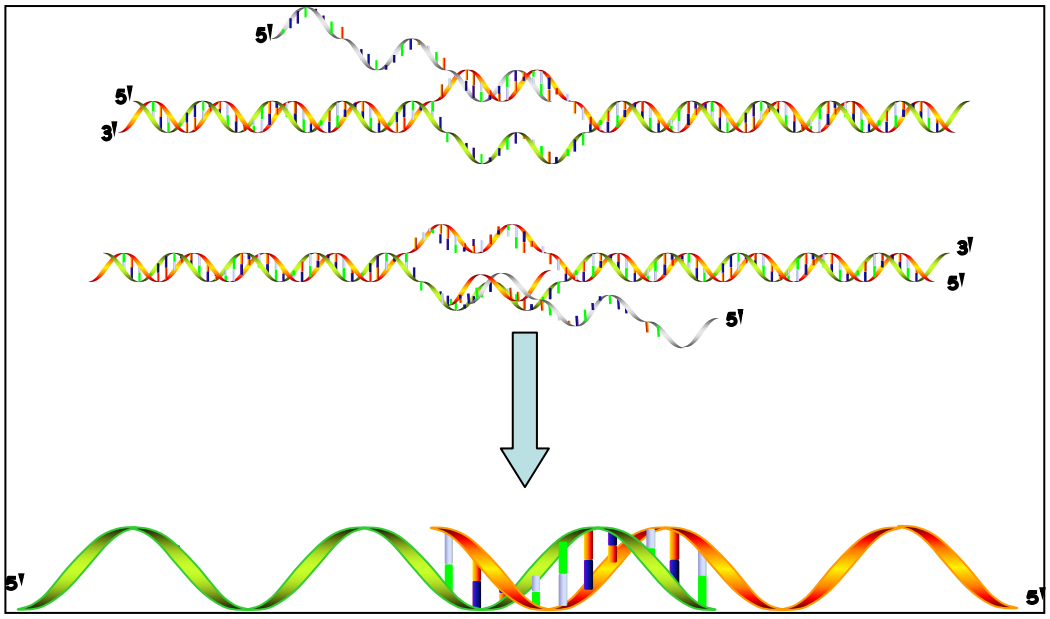


Figure 3. Synthesis of cis natural antisense transcripts (cis-NATs) from both strands of DNA. *Cis* transcripts are transcribed from opposite strands of the same chromosomal position while trans AS transcripts arise from separate loci with complementary sequences. It is thought that the cis transcripts contain longer and more complementary sequences for sense transcripts because of this distinction (11).

AS transcripts are predicted to derive many of their regulatory properties through RNA associated functions; these mechanisms include RNA editing, the masking of RNA elements

used in localization and transport, and RNA interference (11). RNA interference (RNAi) is activated when sense and antisense RNA bind to form double-stranded (dsRNA) pairs in a cell. The dsRNA then binds to Dicer, a protein complex which cleaves the RNA into its fragments. These fragments are then bound by another protein complex, RISC, undergoing a process in which one of the RNA strands is eliminated while the second strand remains bound to the protein complex. This RISC-bound fragment then serves as a probe to detect mRNA molecules that carry a genetic code that can pair with the RNA fragment on RISC. The mRNA molecules are then bound to the protein complex, cleaved and degraded. When such mRNA molecules disappear, the corresponding gene is silenced as little or no protein of the encoded type is made (5,13).

We conducted an analysis of the HCMV transcriptome by obtaining sequence data on cDNA clones that represent viral genes from infected cells. The library was composed of different temporal classes of the HCMV transcripts: immediate early (IE), early (E), and late (L). This analysis yielded the recognition of noncoding and antisense (AS) transcripts that accumulate during infection and are not currently included on available genomic maps. As is shown in Table 1, the AS pairs were classified into one of four categories: full overlap, intronic, convergent, and divergent. The Trgovcich analysis (21) defined the categories as follows. In a pair of genetic sequences in which one sequence completely contains the other, a full overlap is observed. If one of the genes of the sequence derives its start within the intron of the other member of the pair and then extends beyond the start of the pair, it is considered to be intronic. Sequences that are divergent include transcripts whose sense-antisense pairs overlap in their 5' regions in a head-to-head fashion, while convergent transcripts overlap in a tail-to-tail fashion in their 3' regions. Concordant and discordant temporal association refers to the S-AS pairs having

been isolated from the same or different temporal libraries respectively. Classifying transcripts in this way may shed light on potential functions.

Table 1. Genes for which cis-Natural S-AS pairs were identified and their properties.

Pair	Gene*	# S clones	Library	# AS clones	Library	Predicted temporal association	Predicted classification of AS pairs**
1	**UL12**	1	E	2	L	Discordant	Divergent
2	**UL13**	6	IE, L	1	E	Discordant	Divergent, *Full Overlap*
3	**UL21**	3	IE	2	IE	Concordant	Divergent
4	**UL22A**	2	IE	3	IE	Concordant	Intronic
5	UL23	3	IE,L	1	L	Discordant	Divergent
6	**UL24**	2	L	3	L	Concordant	Divergent, Convergent
7	**UL25**	5	L	1	IE	Discordant	Divergent, Intronic
8	**UL29**	3	IE	1	L	Discordant	Divergent
9	**UL36**	3	IE, E	1	E	Concordant	Divergent, *Full Overlap*
10	**UL37**	5	IE, L	1	E	Discordant	Convergent
11	**UL52**	6	E, L	1	L	Concordant	Divergent
12	**UL61**	2	L	10	E, L	Concordant	Divergent
13	**UL62**	3	L	12	E, L	Concordant	Convergent, Divergent, Full Overlap
14	**UL63**	19	E, L	3	L	Concordant	Convergent, Divergent, Full Overlap
15	**UL64**	4	L	21	E,L	Concordant	Convergent, Divergent, Full Overlap
16	**UL65**	20	E,L	3	L	Concordant	Convergent, Divergent, Full Overlap
17	UL66	1	L	18	E,L	Concordant	Divergent, Full Overlap
18	UL67	1	L	17	E,L	Concordant	Divergent, Full Overlap
19	**UL70**	1	L	4	IE, L	Concordant	Divergent, *Full Overlap*
20	**UL72**	1	L	3	E,L	Concordant	Divergent, *Full Overlap*
21	UL73	3	E,L	1	L	Concordant	Full Overlap
22	**UL74**	1	L	3	E,L	Concordant	*Full Overlap*, Intronic
23	**UL88**	1	L	2	L	Concordant	Divergent
24	UL89	1	L	2	L	Concordant	Divergent, Convergent
25	UL92	1	L	1	L	Concordant	Divergent, Intronic
26	**UL102**	2	L	2	L	Concordant	Convergent, Divergent,
27	UL108	2	IE, L	1	IE	Concordant	Convergent
28	UL111	6	IE, L	2	IE, L	Concordant	Convergent, Full Overlap
29	**UL111A**	1	L	2	L	Concordant	Full Overlap
30	UL112	5	IE, L	1	L	Concordant	Divergent
31	UL115	4	IE, L	2	L	Concordant	Divergent, Intronic, *Full Overlap*
32	UL116	2	IE	2	L	Discordant	Divergent, Intronic, *Full Overlap*
33	**RL2**	6	E, L	15	E,L	Concordant	Convergent, Divergent, Full Overlap
34	RL3	7	E, L	189	E,L	Concordant	Convergent, Divergent, Full Overlap
35	RL4	167	E, L	4	E, L	Concordant	Convergent
36	**TRS1**	1	IE, E, L	1	L	Concordant	*Full Overlap*
37	**US32**	1	L	1	L	Concordant	Divergent
38	**US33**	1	L	4	IE, E, L	Concordant	Convergent, Divergent, Intronic

* Bold indicates sequence analysis confirmed a genuine poly (A) tail of the AS member of the pair.
** Italicized 'full overlap' designations are tentative because the 3' ends were not sequenced.
Reprinted with permission of author, (ref.)

CpG Island Methylation. Natural AS transcripts are implicated in gene silencing through methylation of CpG islands associated with gene promoter regions (11). CpG islands are typically greater than 200bp with guanine + cytosine contents greater than 50% and can act as regulatory elements of associated gene functions (19). DNA methylation acts as an epigenetic mechanism that regulates chromosomal stability and gene expression (12). More specifically, methylation of the cytosines in CpG dinucleotides within a promoter-related CpG island can lead to silencing of the gene. The addition of methyl groups to the DNA backbone alters the appearance and structure of the DNA, interfering with transcription factor binding or inducing a repressive chromatin structure, both mechanisms resultant in transcriptional repression (20). This methylation process is known to be involved in genomic imprinting and silencing of genes on the inactive X chromosome. Though cytosine methylation is required for mammalian development (20), it has been shown that exposure to a wide variety of xenobiotics during critical periods of this development can persistently alter the pattern of DNA methylation, resulting in potentially adverse biological effects such as aberrant gene expression (12). Aberrant methylation of cytosines has been associated with certain disease states including cancer. Methylation may affect tumor development by impacting control of cell cycle progression and lead to inactivation of tumor-suppressor genes (1, 20).

Specific goals. The hypothesis that was being tested in this project was that AS transcripts mediate the silencing of viral gene promoters associated with CpG islands and that this is an important mechanism in regulation of the pattern of gene expression.

Materials and Methods.

Cell line and virus. Normal human lung fetal fibroblasts (MRC-5s) were used in this study. MRC-5 cells were maintained in modified Eagle's medium supplemented with 10% fetal calf serum, 1.7mM sodium bicarbonate, 1.4mM sodium chloride, essential and nonessential amino acids, vitamins, and sodium pyruvate at manufacturer-recommended concentrations, (Sigma-Aldrich, InVitrogen). HCMV strain AD169 was obtained from ATCC and was propagated and titrated on MRC-5 cells by plaque assay.

Viral DNA isolation. Confluent MRC-5s in T-150s were exposed to the AD169 strain of HCMV at an MOI of 1.0. Mature virus in the cytoplasmic compartment of cells and in the supernatant were isolated separately at 72 h and 96h postinfection (p.i.). Cell-associated virus was isolated by pelleting cells by centrifugation. The cell pellets were resuspended in 5 ml of 150 mM NaCl, 10mM Tris (pH 7.4), and 1.5 mM $MgCl_2$. After incubation on ice (10 min), NP-40 was added to a final concentration of 0.1%. The lysate was centrifuged at 3,700 rpm for 20 min using a Beckman GS-6R centrifuge. The supernatant was collected and brought to a final concentration of 0.2% dodecyl sulfate (SDS), 0.5 mM EDTA, and 50 mM beta-mercaptoethanol. After incubation on ice and extraction with phenol-cholorofrom, the genomic DNA was precipitated with ethanol, resuspended in 1 mL of Tris-EDTA buffer, and treated with RNase (Sigma-Aldrich). The genomic DNA was further purified by centrifugation in a linear 5 to 20% (wt/vol) potassium acetate gradient at 40,000 rpm for 3.5 h at 20˚C in a Beckman L7 Ultracentrifuge SW60 rotor. Following centriguation, the DNA was collected, precipitated with ethanol, and resuspended in 50ul TE buffer. The mature virions in the media were obtained by transferring the media to separate Falcon centrifuge tubes. The media was spun at 3,000 rpm 4˚C for 20 min in a table top centrifuge. The resultant supernatant was spun again at 15,000 rpm 4˚C

for 1 h. The pellets containing virus were then rinsed 2 X 10mM NaCl 10mM Tris (pH 8.0) and resuspended in 0.5mL proteinase K digestion buffer. Proteinase K was then added to 0.2mg/mL and digested overnight at 37˚C. Phenol-chloroform extractions were performed followed by the addition of 3 volumes ice cold ethanol as well as 10ug carrier RNA. The samples were then precipitated overnight at -20˚C and resusupended in TE buffer. After incubation at room temperature the cells were loaded onto a potassium acetate gradient (5-20%) and run at 40K for 3.5h. Leaving 1-2mLs supernatant in the tubes, the cells were once again ethanol precipitated overnight and resuspended in TE buffer.

Southern Blot Hybridizations. The viral DNA was subjected to restriction digests using the following enzymes; AgeI, BamHI, EagI, HindIII, NheI (New England Biolabs), and BamHI, MluI (Invitrogen). Three reactions using 1µg DNA per reaction were set up for each selected region under study: one double digest including both enzymes, and two single digests, one for the methylation sensitive enzyme and one for the non-methylation sensitive enzyme (methylation sensitive and resistant enzymes for each region are shown in Table 3). The restriction digest reactions were then subjected to agarose gel (0.9% wt/vol) electrophoresis. Nucleic acids were subsequently transferred to positively charged nylon membranes (Millipore) using a Posi-Blot 30-30 pressure blotter (Stratagene). Primers specific for each gene region were used to generate probes as shown in Table 2. All probes were then labeled using a PCR DIG probe synthesis kit (Roche Applied Science) according to manufacturer's instructions. Each probe was selected to correspond to specific nucleotides of the AD169 genome that included a portion of the selected CpG islands under investigation for methylation. These corresponding regions are detailed in Table 4. Hybridization reactions were conducted according to manufacturer's instructions (Roche) and the blots were subjected to a chemiluminescent detection system (DIG High Prime

DNA Labeling Detection Starter Kit II- Ready to use CSPD; Roche). Hybridization reactions were then visualized by exposure to film.

Table 2. Primers for Probe Generation

Primer	Sequence	Genomic Binding Site	PCR product
pL1016-mF	GGC CCT AGA AAG ACC AAG A	19299-19317	probe will recognize between 19299-20369
pL1016-mR	GTT GAA TTT CTT GTC TCC TCT CCC	20392-20369	
pL811-mF	GGG AGT TAT CTT TCT GTC TGG AGG	8641-8618	probe will recognize between 8641-8024
pL811-mR	GAC AGG GAG ATG GAC AGT GTC A	8003-8024	
pL6311-mF	GGA AGG CAT CCA CTT GTA AGG AAC	106329-106352	probe will recognize between 106329-106659
pL6311-mR	GTT CCA CTC CAT CAC ACC ACC	106679-106659	
pL449-mF	CAG GCA GAA TTC CTT CAC TG	164760-164779	probe will recognize between 164760-165780
pL449-mR	GAA CTG TGG TTC CCC CTC TA	165799-165780	
pL715-mF	CAT AGC CCT CGT CCG AGA TGA G	29540-29561	probe will recognize between 29540-30000
pL715-mR	CCT GCA TAC ATG GAG GAG ACC	30020-30000	

Liquid Chromatography and Mass Spectroscopy (LC-MS) . In collaboration with the laboratory of Dr. Kari Green-Church (The Ohio State University) genome-wide CpG island methylation (5-methyl deoxycytosine) will be measured by reversed-phase high-pressure liquid chromatography and mass spectroscopy analysis. Through this process, using a modified version of the method developed by Kuo et. al (10), the level of cytosine methylation will be measured by chromatographic separation and mass spectroscopy analysis of deoxyribonucleotides. Specifically, we will enzymatically digest viral DNA to release deoxribonucleotides and then

measure 5-methyl-2'deoxcytidine-5'-monophosphate (5mdCMP) as a proportion of total cytosines (5mdCMP + dCMP). To refine the methodology and calibrate the instruments the following standards (Sigma) were used and prepared according to manufacturer's instructions: 2'Deoxycytidine 5'-monophosphate (dCMP), 5-Methyl-2'-deoxycitidine, 5'-Monophosphate (5mdCMP), 2'-Deoxyguanosine 5'-monophosphate (dGMP), Thymidine 5'-monophosphate (TMP), and 2'-Deoxyadenosine 5'-monophosphate(dAMP). Test samples were derived from viral DNA. Viral DNA was isolated from infected cells and subjected to enzymatic hydrolysis to ensure removal of RNA that could result in inaccurate measurement of the constituent nucleotides in the sample. This was accomplished by dissolving 50ug DNA into 300ul of 1X TE followed by the addition of RNAse A (final conc. 100ug/mL) and RNAse T1 (final conc. 2,000units/mL), incubating the mixture at 37˚C for 2h. Following incubation, an equal volume of phenol/chloroform/isoamyl alcohol was added and centrifuged for 2min at 15,000rpm in a table-top microcentrifuge. The aqueous layer was then collected in a clean tube and the DNA was precipitated using 0.1 vol of 3M sodium acetate and 2.5 vol of absolute ethanol. The DNA was then recovered by centrifugation and removal of the ethanol supernatant. After rinsing the DNA pellet with 70% ethanol, it was resuspended in 100ul TE. DNAse I (final conc. 50 ug/mL) was then added and incubated at 37˚C for 14 hr. DNAse is an endonuclease that catalyzes the endonucleolytic cleavage of double and single stranded DNA to yield 5'-phosphodinucleotide and 5'-phosphooligonucleotide end-products. The product of this hydrolysis is a complex mixture of 5'-phosphate mononucleotides and oligonucleotides. Next, Benzonase nuclease (final conc. 50ug/mL) was added followed by a 7 hr. incubation period at 37˚C. Benzonase is an endonuclease that hydrolyzes double- and single-stranded DNA and RNA yielding 5'-phosphomononucleotide and 5'-phosphooligonucleotide end-products. Finally, digestion was

completed with nuclease P1 that also yields 5'-phosphomononucleotide and 5'-phosphooligonucleotide end-products.

The following steps will be performed at the Campus Mass Spectrometry and Proteomics Facility. The details of these protocols were provided to us by Dr. Kari Green-Church, director of the facility. Samples will be prepared in an acidified liquid solution for chromatographic separation. Upon injection onto a silica column, the nucleotide mixture will be separated according to its solubility in a polar solution. Elutants will be transferred into an electrospray ionization source, a metal capillary charged to a potential of roughly 3,000V. The liquid chromatographic/autosampler system will consist of a Micro UltiMate™ 3000 system from LC-Packings A Dionex Co (Sunnyvale, CA). Eluent from the column will be directly injected into a Micromass Qq-TOF™ II (Micromass, Wythenshawe, UK) mass spectrometer equipped with an orthogonal electrospray source (Z-spray) operated in positive ion mode. Sodium iodide will be used for mass calibration within a calibration range of a mass-to-charge ratio (m/z) of 100 - 2000. The optimal ESI conditions to employ will include: capillary voltage 3000 V, source temperature 110^{o}C and a cone voltage of 55 V. The ESI gas will be nitrogen. Q1 will be set to optimally pass ions from m/z 100 – 2000 and all ions transmitted into the pusher region of the TOF analyzer will be scanned over the mass range of the samples, m/z 100-1000, with an integration time of 1sec. The eleutant liquid is sprayed out of the capillary to form an aerosol and the voltage produces highly charged droplets that become smaller and smaller until Coulomb fission occurs, producing a gas phase charged analyte ion. Coulomb fission is a two step process in which multiple Coulomb excitations of an actinide target nucleus by a heavy projectile occur followed by radioactive decay (14). Based on the masses of the resultant ions, we will be able to

measure the relative amounts of each nucleotide in the mixture as well as measure the ratio of 5mdCMP and dCMP in our viral DNA sample.

Results

Mapping CpG Islands.

To verify the presence and identify the exact location of CpG islands in the HCMV genome, a map of putative islands was developed using a program from the European Bioinformatics Institute. The program was able to calculate the number of cytosine-guanine dinucleotides in comparison to the total amount of nucleotides in the genome. The data is shown in Figure 4. In collaboration with Dr. Ilya Ioshikhes (Department of Bioinformatics, The Ohio State University) we generated a map of the CpG islands over the length of the HCMV genome (strain AD 169) and determined whether the CpG islands were promoter related or promoter unrelated based on our transcriptome study (Table 1) and an algorithm developed by Dr. Isohikhes. We found that out of the 206 CpG islands found in the HCMV genome, 65 are promoter related and 59 associate with a known gene, (the remaining 6 suggest putative unmapped genes may exist). CpG islands between 500bp upstream or 1500 bp downstream of a known or predicted transcriptional start site were considered as promoter related (9).

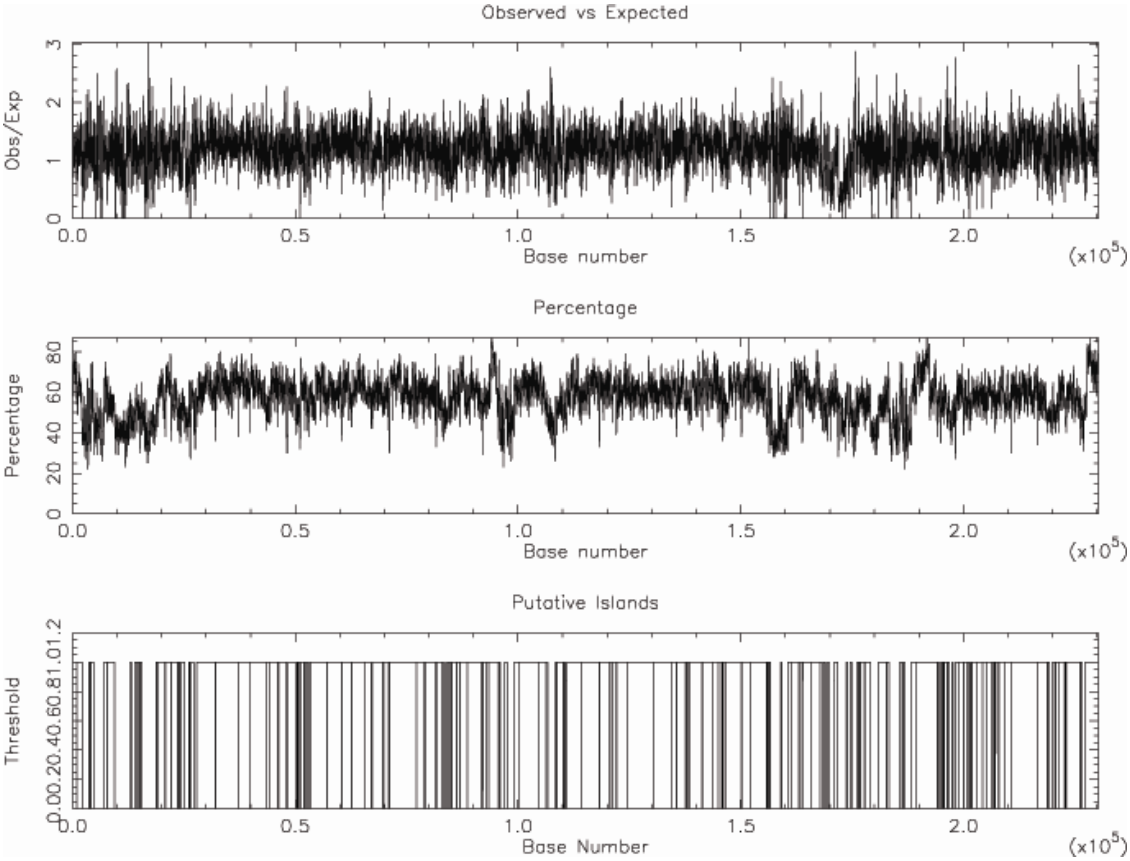


Figure 4. CpG island analysis of the HCMV genome. Portions of the genome (at least 200bp) that possess guanine + cytosine contents greater than 50% are known as CpG islands. These islands can act as regulatory elements of associated gene functions (19). Using a program developed by Gardiner-Garden and Frommer (6) the CpG plot of the HCMV genome was analyzed to produce a parameter report. Using default settings provided by the program, the observed vs. expected diagram depicts the possibility for a CpG island where the minimum average observed to expected ratio of C plus G to CpG is above 0.6 within a window of 1,000 base pairs. The minimum percentage diagram establishes the minimum average percentage of G plus C (50% in this case) that is required before a CpG island is detected within a window of 1,000 base pairs. To be reported, each CpG island was required to be at a minimum of 200 base pairs. The resultant putative islands were then plotted, demonstrating the rich GC composition of the HCMV genome.

Identification and Selection of AS cDNA clones overlapping CpG Islands

We next linked the CpG islands to transcripts identified in our study (21), focusing on AS transcripts that overlap the CpG islands. We focused our attention to those CpG islands with overlapping AS transcripts that fell into the late temporal class of viral genes. Although some promoter elements are known to generally correlate with immediate early (IE) and early (E) genes, no clear viral promoter elements correlate with expression of late genes. We therefore

speculated that perhaps epigenetic mechanisms contribute to or even dominate regulation of late genes. Viral genes are expressed in a temporal cascade and are classified as immediate early, early, or late, late genes being expressed only after replication of the viral DNA genome. In predicting that CpG island methylation acts to silence these late genes at early stages of infection and that DNA synthesis acts to remove this block to gene expression, only genes located in the late temporal library were chosen for the experiment.

Once the positions of the overlapping transcript (clone) and CpG islands were determined, restriction enzymes were chosen based on inclusion of their cut sites in the gene regions. For each region, a methylation sensitive and a methylation insensitive enzyme were selected as shown in Table 3. It was expected that in methylated regions of the genome, sensitive enzymes would be unable to "cut" the DNA, whereas the regions that did not experience methylation would be digested to produce DNA fragments of predicted sizes.

Table 3. Antisense Transcripts and Restriction Enzymes that overlap late viral genes and map to CpG Islands.

Gene Region harboring CpG island	Clones overlapping this CpG island in AS orientation	Orientation of transcript relative to gene (s)	Promoter Relation	CpG position	Enzymes Utilized	Methylation sensitive
UL12	pL1016	AS/S	+	19061 - 20772	BamHI/EagI	No/Yes
RL11	pL811	AS/S	+	7732- 9437	BamHI/MluI	No/Yes
UL115	pL449	AS	+	163880 - 165863	BamHI/AgeI	No/Yes
UL72	pL6311	AS/S	-	106316 - 106680	BamHI/NheI	No/Yes
UL24	pL715	AS	-	27837- 32047	HindIII/EagI	No/Yes

Table 4. Probes used for Southern blot analyses.

Gene Region harboring CpG island	Clones overlapping this CpG island in AS orientation	Gene Position	CpG position	Probe
UL12	pL1016	19140-20843	19061 - 20772	19299-20369
RL11	pL811	8726-9430	7732-9437	8024-8641
UL115	pL449	164758-166039	163880 - 165863	164760- 165780
UL72	pL6311	106326-108586	106316 - 106680	106352-106659
UL24	pL715	28849-29720	27837-32047	29540-30000

Southern blot analysis of methylation in the HCMV genome. Once overlap was observed between the CpG islands and AS transcripts in the genome, the next step of our experiment was to explore whether or not methylation occurs in these regions. Restriction sites associated with each CpG island were mapped. The sites associated with the RL11 CpG island are shown as an example in Figures 5A-D.

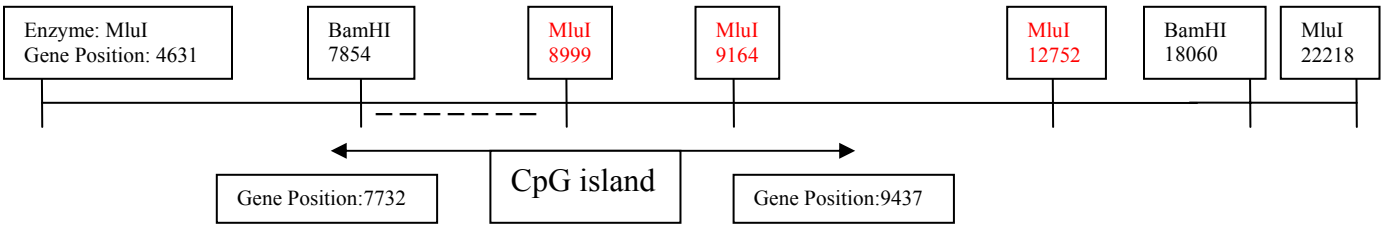
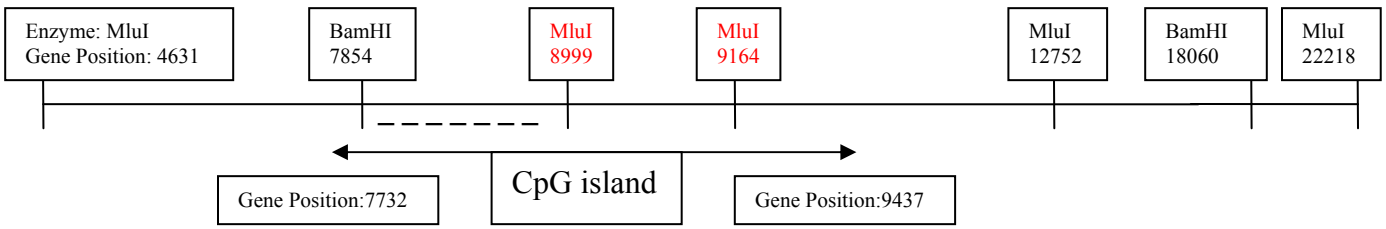
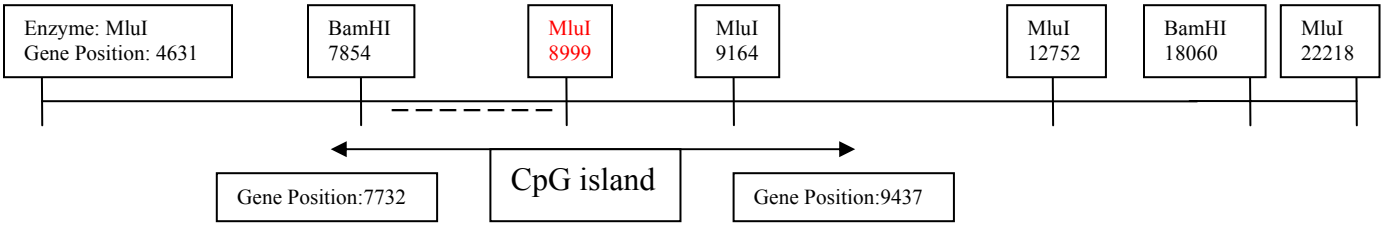


Figure 5 A. In a double digest using BamHI (the methylation resistant enzyme) and MluI (the methylation sensitive enzyme) a 10kb fragment is produced when methylation occurs in the MluI sites marked in red. The double-headed arrow represents the CpG island. The dashed line represents the probe.



B. In a double digest using BamHI (the methylation resistant enzyme) and MluI (the methylation sensitive enzyme) a 4.8kb fragment is produced when methylation occurs in the MluI sites marked in red. The double-headed arrow represents the CpG island. The dashed line represents the probe.



C. In a double digest using BamHI (the methylase resistant enzyme) and MluI (the methylation sensitive enzyme) a 1.3kb fragment is produced when methylation occurs in the MluI sites marked in red. The double-headed arrow represents the CpG island. The dashed line represents the probe.

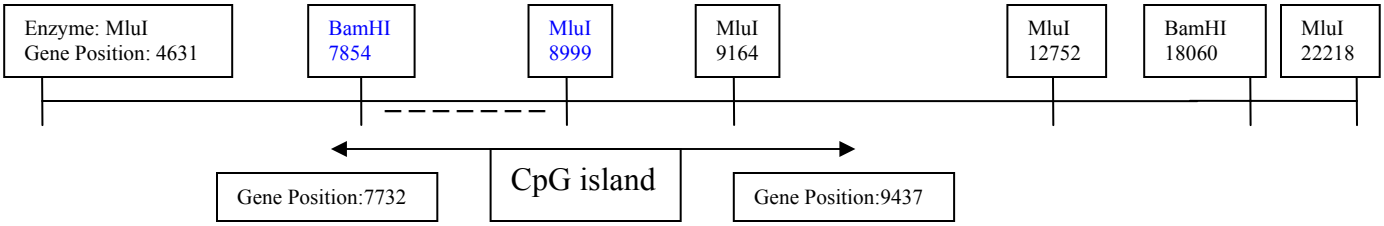


Figure 5D. In a double digest using BamHI (the methylase resistant enzyme) and MluI (the methylation sensitive enzyme) a 1.1kb fragment is produced when digestion occurs at the cut sites marked in blue, methylation being absent from the MluI sites. The double-headed arrow represents the CpG island. The dashed line represents the probe.

Banding patterns were then predicted for each region, acknowledging the variable fragment sizes that are generated when the sites are methylated as compared to not methylated. These predictions were made assuming that when methylation sensitive enzyme sites are methylated, digestion by the enzyme will be inhibited, producing fragment sizes that differ from the banding pattern generated when the enzyme is active (not affected by methylation) and able to cut the DNA. These expected banding patterns are illustrated in Table 5.

Viral DNA from mature virions was then isolated and subjected to restriction digests. These digests were separated by agarose gel electrophoresis, and upon transfer to nylon membranes, reacted with the probes shown in Table 4. In completing the first Southern blot, the

obtained data was inconsistent with methylation of the restriction enzyme recognition sequences utilized in the experiments as is seen in Figure 6 and Table 5.

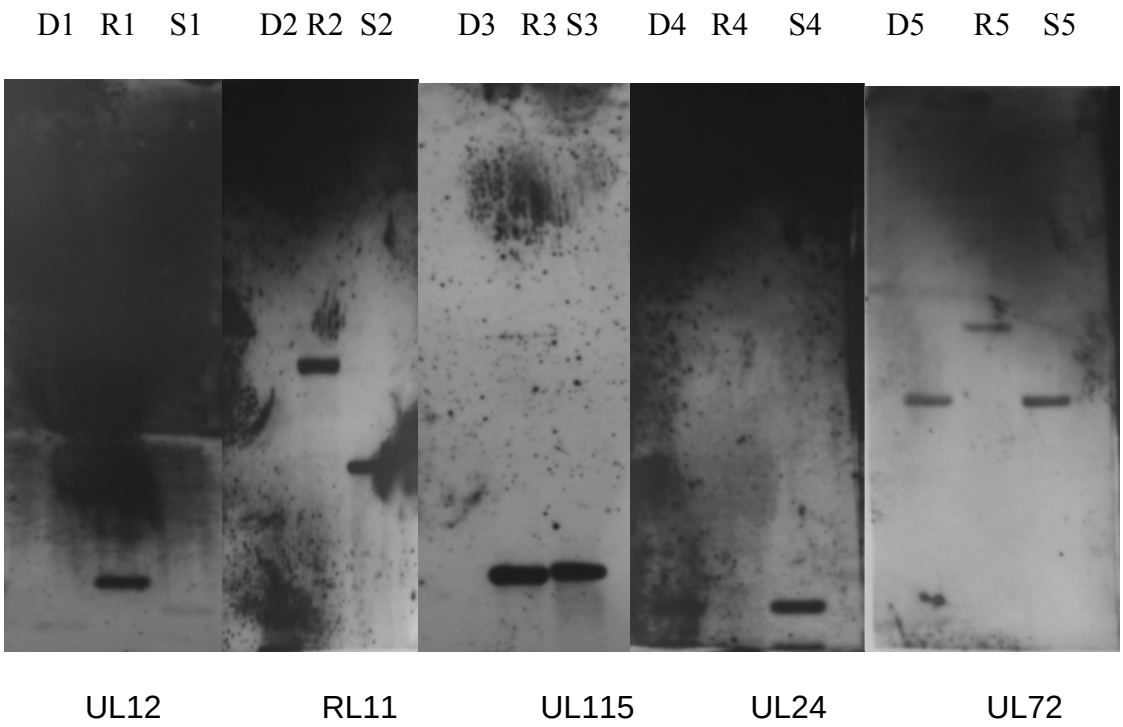


Figure 6. Southern Blot #1. D:double digest; R: methylation resistant enzyme; S: methylation sensitive enzyme. One microgram of MRC-5 cytoplasmic DNA was electrophoresed after digestion with the following enzymes. BamHI+EagI (D1); BamHI (R1); EagI (S1); BamHI+MluI (D2); BamHI (R2); MluI (S2); BamHI+AgeI (D3); BamHI (R3); AgeI (S3); HindIII + EagI (D4); HindIII (R4); EagI (S4); BamHI+ NheI (D5); BamHI (R5); NheI (S5). After electrophoresis the fragments were transferred to nylon and the nylon was hybridized with probes specific for each gene region, (Table 4). Exposure time was 20 minutes. UL12, RL11, UL24, and UL72 produced double digest fragments that corresponded with measurements predicted to indicate the absence of methylation in the genome. Although UL115 did not produce fragments in its double digest that correlated with those expected for negative methylation, this was likely due to the fact that the expected negative fragments, 867bp and 841bp, would have run off of the gel as indicated by the banding patterns that were displayed.

Table 5. Results for methylation in selected HCMV gene regions based on Southern blot experiment #1.

Gene Region harboring CpG island	Clones overlapping this CpG island in AS orientation	Predicted sizes of fragments - CH_3	Predicted sizes of fragments + CH_3	*Estimated sizes of fragments generated by double digests	Evidence for possible methylation
UL12	pL1016	1877bp	167bp, 228 bp, 488bp, 1389bp	1.3kbp	-
RL11	pL811	10+kb, 4898bp, 1310bp	1145bp	1.1kbp	-
UL115	pL449	2110bp, 1708bp	867bp, 841 bp	N/A	?
UL72	pL6311	14kb, 9529bp, 7286bp	6977bp, 309bp	6.9kbp or 7.2kbp?	?
UL24	pL715	3854bp, 2742bp, 1210bp	1506bp, 1183bp, 27bp	1.6kbp 1.2kbp	-

*At the time of the experiment, the appropriate ladder was not available in the laboratory to serve as a comparison, so fragment sizes (Kbp) were roughly estimated.

To confirm this result, the experiment was repeated. Although these studies were inconclusive due to DNA degradation (see appendix), one of the experiments suggested possible methylation in the RL11 CpG island. To further explore this, we repeated the restriction analysis of this single gene region using DNA preparations generated from different time points of the same infection (14h, 44h, 72h, and 96h). The data from the 72h and 96h time points are shown in Figure 7. The banding pattern was not consistent with methylation at these or the other time points (data not shown). Based on predictions outlined in Fig. 5D and the results summarized in Table 6, we conclude the Mlu I sites are not methylated in the RL11-associated CpG island.

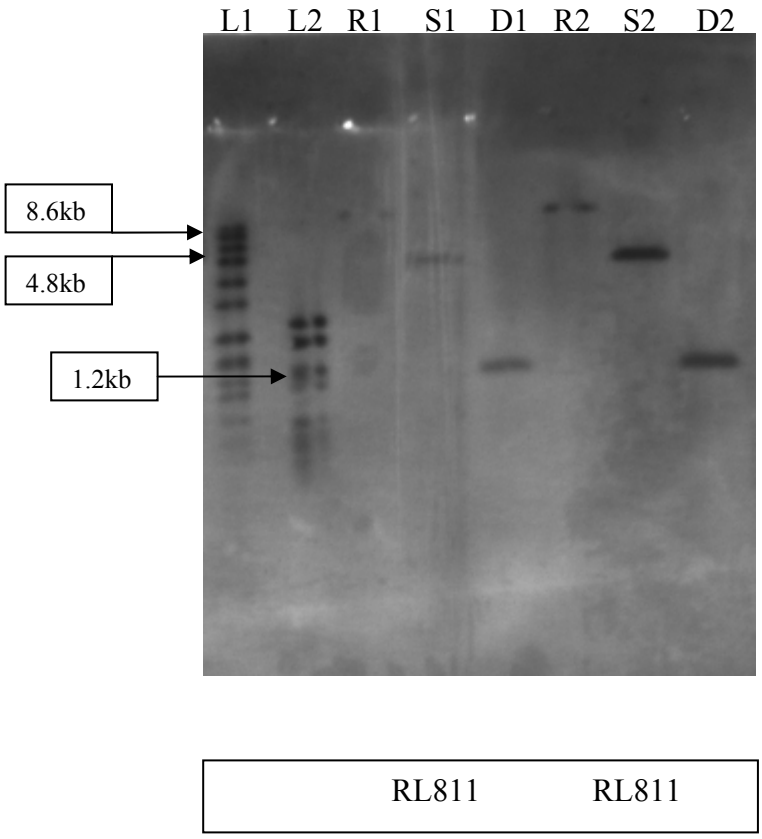


Figure 7. Southern blot #5. L1: DIG-labeled molecule weight marker VII (Roche); L2 DIG-labeled molecule weight marker VI (Roche); D: double digest; R: methylation resistant enzyme; S: methylation sensitive enzyme. One microgram of MRC-5 supernatant [72h (S1, R1, D1) 96h (S2, R2, D2)] DNA was electrophoresed after digestion with the following enzymes: BamHI (R1,R2); MluI(S1,S2); BamHI+MluI (D1,D2). After electrophoresis, the gel was blotted and the filter was hybridized with probes specific for each gene region, (Table 4). Exposure time was 20 minutes.

Table 6. Results for methylation in selected HCMV gene regions based on Southern blot experiment #5.

Gene Region harboring CpG island	Clones overlapping this CpG island in AS orientation	DNA preparation from MRC-5 cells	Predicted sizes of fragments - CH_3	Predicted sizes of fragments + CH_3	*Estimated sizes of fragments generated by double digests	Evidence for possible methylation
RL11	pL811	72h	10+kb 4898bp 1310bp	1145bp	1.1kbp	-
RL11	pL811	96h	10+kb 4898bp 1310bp	1145bp	1.1kbp	-

We conclude from this series of experiments that the CpG dinucleotides are not methylated at the specific restriction sites tested. The major limitation of this Southern blot technique was the limited number of CpG dinucleotides within the genome that could be analyzed. Therefore, concurrent with these studies, we were developing methods to analyze methylation using global approaches. It is with these additional analyses that we will be able to make a concrete conclusion on CpG island methylation status in the genome and its subsequent correlation or lack of correlation with AS transcripts. Until that time, our present conclusion cannot be extended to the entire genome.

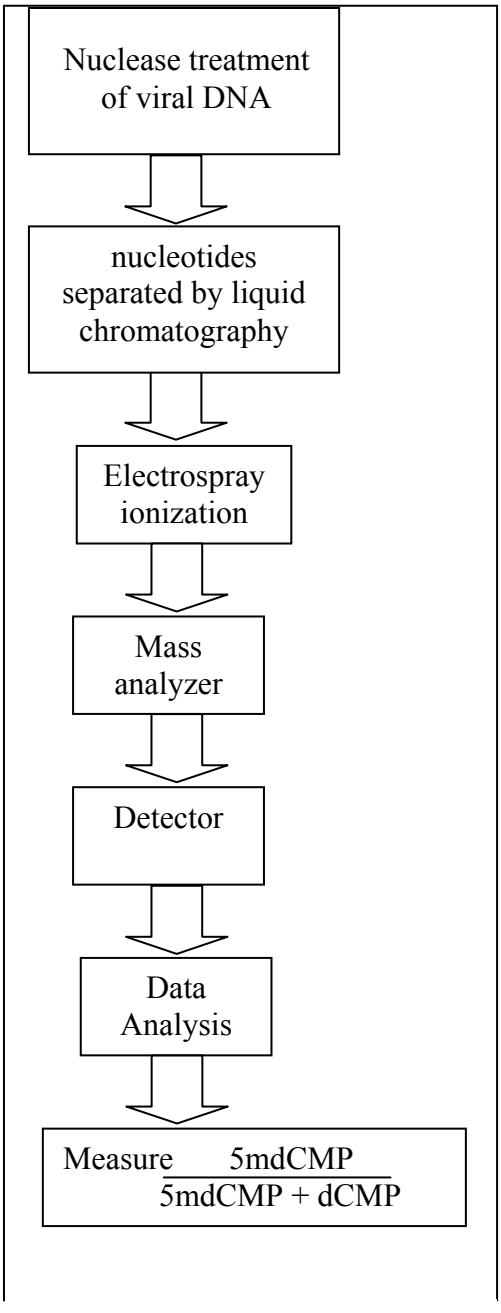


Figure 8. General scheme of liquid chromatography-mass spectroscopy studies.

Global Methylation Studies: LC-MS Analysis of nucleotides derived from mature virus.

An overview of the LC-MS approach is shown in Fig. 8.Viral DNA was hydrolyzed according to the method developed by Kuo et. al (Materials and Methods) and the mass spectrometer was calibrated using standard nucleotide solutions (Sigma) to generate the data shown in Figure 9. The methods developed were successful in separating and identifying A, C, G and T nucleotides by LC-MS. The small extra peak observed in the 2-deoxyguanidin monophosphate standard is an impurity. Unfortunately, the methylated standard purchased from Sigma was too impure to provide a reliable standard. However, because we expect 5mdCMP to have both a

different retention time and different mass than dCMP, we predict based on the standard analysis that we will be able to measure the amount of methylated cytosines in the viral genome. At the writing of this paper, we are awaiting results from the LC-MS analysis of the viral DNA . If HPLC-MS is successful in producing evidence of methylated cytosines in the genome, we will attempt to use this approach to measure changes in the methylation of viral DNA over time in infected cells. Localization of the methylated cytosines will be determined by ChIP-on-chip methodology (Future Studies).

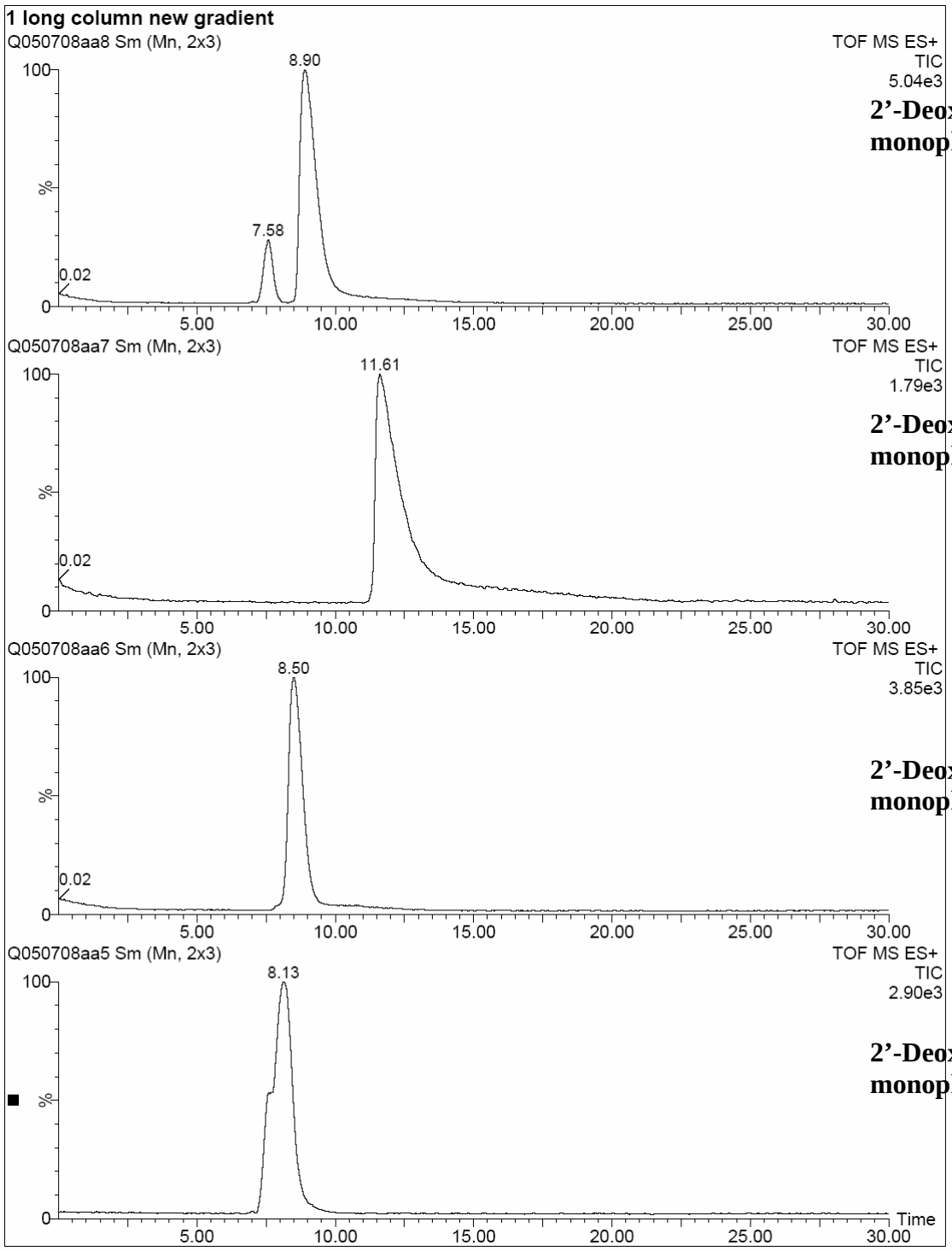


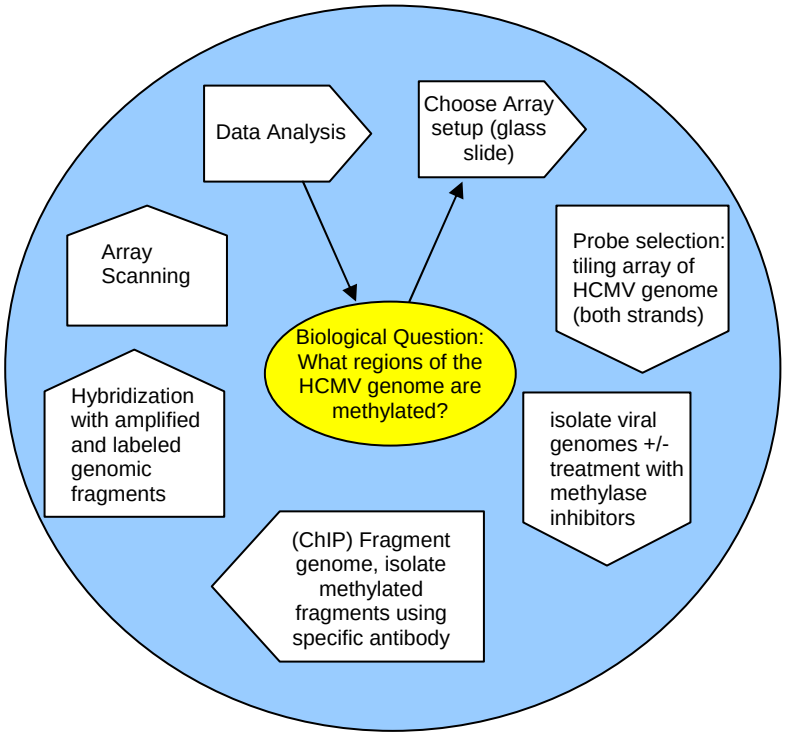
Figure 9**.** Mass Spectroscopy Chromatograph of Nucleotide Standards. Chromatograph generated from the Sigma nucleotide standards used to optimize methods and calibrate the mass spectrometer for the experiment. Each separate nucleotide produces its own distinct peak so that we will be able to determine the concentration of each within the HCMV genome. Methylated cytosine (not shown due to problems with the standard) will appear as its own peak, separate from unmethylated cytosine.

Future Studies

Immunoprecipitation and Gene Array. If we find methylated cytosine residues by LC-mass spectroscopy analysis we will next quantitate the abundance of methylated cytosines at different phases of infection. However, this anlaysis gives no information regarding the location of the methylated cytosines. To answer this question, we will collaborate with Dr. Peter Charles (University of North Carolina) to perform a second genome-wide methylation analysis using a chIP-on-chip procedure that combines a modified chromatin immunoprecipitation with glass slide microarray technology. ChIP-on-chip, also known as genome-wide location analysis, is a technique used for the isolation and identification of DNA sequences occupied by specific DNA binding proteins in a cell that allows us to analyze how regulatory proteins interact with the genome of living cells. These identified protein binding sites may indicate functions of transcriptional regulators including promoters, enhancers, repressor insulators, and boundary elements/sequences. In our case, we will not use antibodies against DNA binding proteins. Rather, we will use an antibody that specifically recognizes methylated cytosine residues in the DNA. The overview of the approach is illustrated in Figure 10. To begin, Dr. Charles developed a tiling array of the entire HCMV genome (GenBank accession no. NC_001347); in which each probe represents 50bp of the genome and each set of adjacent probes overlaps the previous by 25bp. Next, isolated viral DNA from mature virions will be fragmented using a combination of restriction enzymes. These fragments will be reacted with antibody specific for 5-methylcytosine. Fragments isolated in this approach will be amplified using labeled (cyanine 5) oligonucleotides and reacted with the glass slides. As a positive control, total fragmented viral DNA will be amplified and hybridized to the slide. As a negative control, viral DNA will be reacted with a methylase enzyme to remove any methyl groups on the cytosines prior to the

immonoprecipitation step. We expect these studies to identify the genomic location of methylated cytosine residues. If we identify methylated regions, we will then determine whether they correlate with (1) known or predicted promoter regions, (2) antisense transcripts, and (3) temporal regulation of associated transcripts.

Figure 10. ChIP-on chip methodology overview



Discussion

In this study, we isolated viral DNA and subjected it to restriction digests and Southern blot analyses in an attempt to determine if antisense transcripts correlate with methylation of viral DNA at specific promoter related and unrelated CpG islands. The major assumption being that the hypermethylation of these cytosines in CpG dinucleotides in the HCMV genome may provide an important mechanism in regulating the pattern of viral gene expression through silencing of CpG island promoter elements. Our hypothesis was refuted by our data, methylation sensitive enzymes remaining active in selected gene regions, indicating the absence of methylation. Though theses initial studies do not support methylation at the specific genomic sites analyzed, two subsequent experiments are currently being conducted in order to test this hypothesis across the entire genome. These experiments, which include liquid chromatography in combination with mass spectroscopy and chromatin immunoprecipitation in combination with microarray analysis, will provide a more global investigation of the methylation status of the genome. This is important due to the fact that our studies targeted only a few CpG islands, and only a few specific cytosine nucleotides within these islands. To our knowledge, these will be the first studies to systematically analyze viral DNA for methylation across the entire genome. Therefore, both positive and negative findings may reveal new information to help us determine if this virus utilizes cellular methylation machinery to modify its genomic DNA. If methylation is discovered, and is found to correlate with known or predicted promoter elements, this will provide the rationale to pursue studies to determine if methylation is regulated by antisense transcription. It will be through these experiments that we will be able to confirm or refute the hypothesis that AS transcripts influence the life cycle of HCMV by mediating the silencing of viral gene promoters associated with CpG islands.

References

1. **Campanero, M.R., Armstrong, M.I., Flemington, E.K.** 2000. CpG methylation as mechanism for the regulation of E2F activity. PNAS.vol 97.no 12: 6481-6486
2. **Centers for Disease Control and Prevention.** 2006. National Center for Immunization and Respiratory Disease. Atlanta, GA. np.
3. **Ching, K.** 2007. ChIP-on-chip. RenLabs. University of California, San Diego. San Diego, CA. np
4. **Delves, P.J., Martin, S.J., Burton, D.R., and Roitt, I.M.** 2006.Roitt's Essential Immunology, 11th edition. Blackwell Publishing Inc. Malden, MA.
5. **Fire A., Xu S.Q., Montgomery M.K., Kostas S.A., Driver S.E., Mello C.C**. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 1998; 391: 806-811.
6. **Gardiner-Garden, M. and Frommer, M.** 1987. CpG islands in vertebrate genomes. Journal of Molecular Biology. 196(2): 261-282
7. **Heard, E. and Disteche, C.M.** 2006. Dosage compensation in mammals: fine-tuning the expression of the X chromosome. Genes and Development. 20:1848-1867
8. **Huang, E., and Johnson, R.A**. 2000. Human cytomegalovirus − no longer just a DNA virus. Nature Medicine**.** 6**:** 863 - 864
9. **Isochikhes, I.** 2007. The Ohio State University Dept. of Bioinformatics. Columbus, OH. np
10. **Kuo, K.C., McCune, R.A., and Gehrke, C.W.** 1980. Quantitative reversed-phase high-performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. Nucleic Acids Res. **8**: 4763-4776
11. **Lavorgna, G., D. Dahary, B. Lehner, R. Sorek, C.M. Sanderson, and G. Casari**. 2004. In search of antisense. Trends Biochem. Sci. 29. 88-94
12. **Li, S., Hursting, S.D., Davis, B.J., McLachlan, J.A., and Barrett,** J.C. 2003. Environmental Exposure, DNA Methylation, and Gene Regulation**,**Lessons from Diethylstilbesterol-Induced Cancers. Annals of the New York Academy of Sciences 983:161-169
13. **McVittie, B**. 2006. Nobel Prize for Epigenetics. Epigenome NoE. Tewkesbury Printing Company, UK. np
14. **Oberacker, V.T., Pinkston, W.T., and Kruse, H.G.W.** 1985. Rep. Prog. Phys. **48**:327-374
15. **Pass, R.** 2001. Cytomegalovirus, vol.2. Lippincott, Williams, and Wilkins, Philadelphia, PA.
16. **Sanchez, V. and Spector, D.H.** 2002. Virology. CMV makes a timely exit. Science. 2002 Aug 2;297(5582):778-779.
17. **Soderberg-Naucler, C**. 2006. Does cytomegalovirus play a causative role in the development of various inflammatory diseases and cancer? J. Intern. Med. **259**: 219-246
18. **Somes, N.** 2003. What is CMV? A "BABES" perspective. Seattle Treatment Education Project. STEP Ezine. np.
19**. Takai, D. and Jones, P**. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. PNAS vol. 99 (6): 3740-3745
20. **Weber, M., Davies, J.J.,Wittig, D. Oakeley, E.J., Hasse, M., Lam, W.L., and Schubeler, D.** 2005. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nature Genetics. Vol **37** no 8. 853-862.

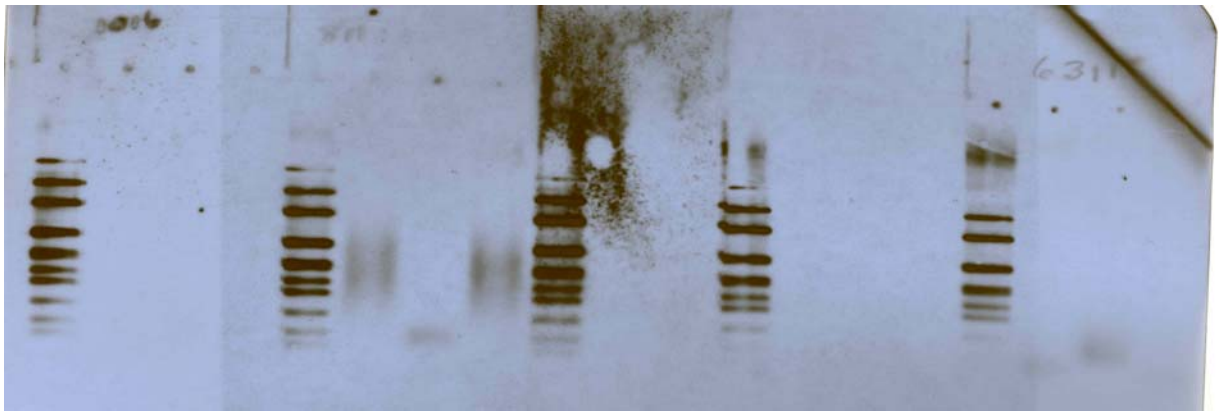
21. **Zhang, G. Raghavan, B. Kotur, M., Cheatham, Jacquelyn, Sedmak D., Cook, C. Waldman, J. and Trgovcich, J.** 2007. Antisense Transcription in the Human Cytomegalovirus Transcriptome. J. Virol. **81:** 11267-11281

Appendix

Five total Southern blots were included in our experiments; however, we chose to report on only two of those blots (experiments 1 and 5) in the main body of this paper. Our research is presented in this format due to the fact that those blots constructed in experiments 2, 3, and 4 were compromised due to DNA degradation and possible probe failure. We present them here to verify their inclusion in our project.

In completing a second Southern blot, data from the first experiment was neither supported nor refuted as smearing and indistinct banding patterns indicated possible DNA degradation or contamination detected by the probes. However, though no concrete refute could be concluded because of the indistinct nature of the fragments, bands in the UL72 region did appear to negate the idea of methylation. It is important to note that while the five minute exposure (Figure A1) is illustrated in these results, subsequent exposures were developed but did not produce any additional fragments. The five minute exposure is shown here as it was the most distinct of the films due to added background in the latter exposures as the membrane began to dry.

L D1 R1 S1 L D2 R2 S2 L D3 R3 S3 L D4 R4 S4 L D5 R5 S5



UL12 RL11 UL115 UL 24 UL72

Figure A1. Southern blot #2. L: DIG-labeled molecule weight marker VII (Roche); D: double digest; R:methylation resistant enzyme; S: methylation sensitive enzyme. One microgram of MRC-5 cytoplasmic DNA was electrophoresed after digestion with the following enzymes: BamHI+EagI (D1); BamHI (R1); EagI (S1); BamHI+MluI(D2); BamHI(R2); MluI (S2); BamHI+AgeI (D3); BamHI (R3); AgeI (S3); HindIII+EagI (D4); HindIII (R4); EagI (S4); BamHI+NheI (D5); BamHI (R5); NheI (S5). After electophoresis, the gel was blotted and the filter was hybridized with probes specific for each gene region, (Table 4). Exposure time was 5 minutes.

Table A1. Results for methylation in selected HCMV gene regions based on Southern blot experiment #2.

Gene Region harboring CpG island	Clones overlapping this CpG island in AS orientation	Predicted sizes of fragments in double digest if methylation sensitive sites are methylated	Predicted sizes of fragments in double digest if methylation sensitive sites are not methylated	*Estimated size of fragments generated by from Double digest	Evidence for possible methylation
UL12	pL1016	1877bp 1389bp	488bp 228bp 167bp	N/A	N/A
RL11	pL811	10+kb 4898bp 1310bp	1145bp	N/A	N/A
UL115	pL449	2110bp 1708bp	867bp 841bp	N/A	N/A
UL72	pL6311	14kb 9529bp 7286bp	6977bp 309bp	N/A	N/A
UL24	pL715	3854bp 2742bp 1210bp	1560bp 1183bp 27bp	N/A	Though not distinct, the low banding patterns observed in this blot appear to refute the possibility of methylation.

Trial three (Figure A2) was also invalid in providing verification of our results, as the DNA was subjected to an inappropriate electrophoresis time, causing the samples to run off of the gel and produce indistinguishable results.

L D1 R1 S1 L D2 R2 S2 L D3 R3 S3 L D4 R4 S4 L D5 R5 S5

UL 12 RL11 UL115 UL24 UL72

Figure A2. Southern blot #3. L: DIG-labeled molecule weight marker VII (Roche); D: double digest; R:methylation resistant enzyme; S: methylation sensitive enzyme. One microgram of MRC-5 cytoplasmic DNA was electrophoresed after digestion with the following enzymes: BamHI+EagI (D1); BamHI (R1); EagI (S1); BamHI+MluI(D2); BamHI(R2); MluI (S2); BamHI+AgeI (D3); BamHI (R3); AgeI (S3); HindIII+EagI (D4); HindIII (R4); EagI (S4); BamHI+NheI (D5); BamHI (R5); NheI (S5). After electophoresis, the gel was blotted and the filter was hybridized with probes specific for each gene region, (Table 4). Exposure time was 30 minutes.

Table A2. Results for methylation in selected HCMV gene regions based on Southern blot experiment #3.

Gene Region harboring CpG island	Clones overlapping this CpG island in AS orientation	Predicted sizes of fragments in double digest if methylation sensitive sites are methylated	Predicted sizes of fragments in double digest if methylation sensitive sites are not methylated	*Estimated size of fragments generated by from Double digest	Evidence for possible methylation
UL12	pL1016	1877bp 1389bp	488bp 228bp 167bp	N/A	N/A
RL11	pL811	10+kb 4898bp 1310bp	1145bp	N/A	N/A
UL115	pL449	2110bp 1708bp	867bp 841bp	N/A	N/A
UL72	pL6311	14kb 9529bp 7286bp	6977bp 309bp	N/A	N/A
UL24	pL715	3854bp 2742bp 1210bp	1560bp 1183bp 27bp	N/A	N/A.

In trial four, three regions were selected for further investigation as initial interpretation of previous experiments evidenced a possibility of methylation in these regions. Subsequent experiments were then begun before a second analysis showed the initial interpretation to be flawed. Of the three regions however, RL11 did appear to produce a positive confirmation of methylation, generating a 10kb fragment that was predicted to occur if all three methylation sensitive enzyme sites in the region were methylated (Figure 5A). Continued smears and indistinct banding in the blot once again indicate DNA degradation and/or complications with the probes.

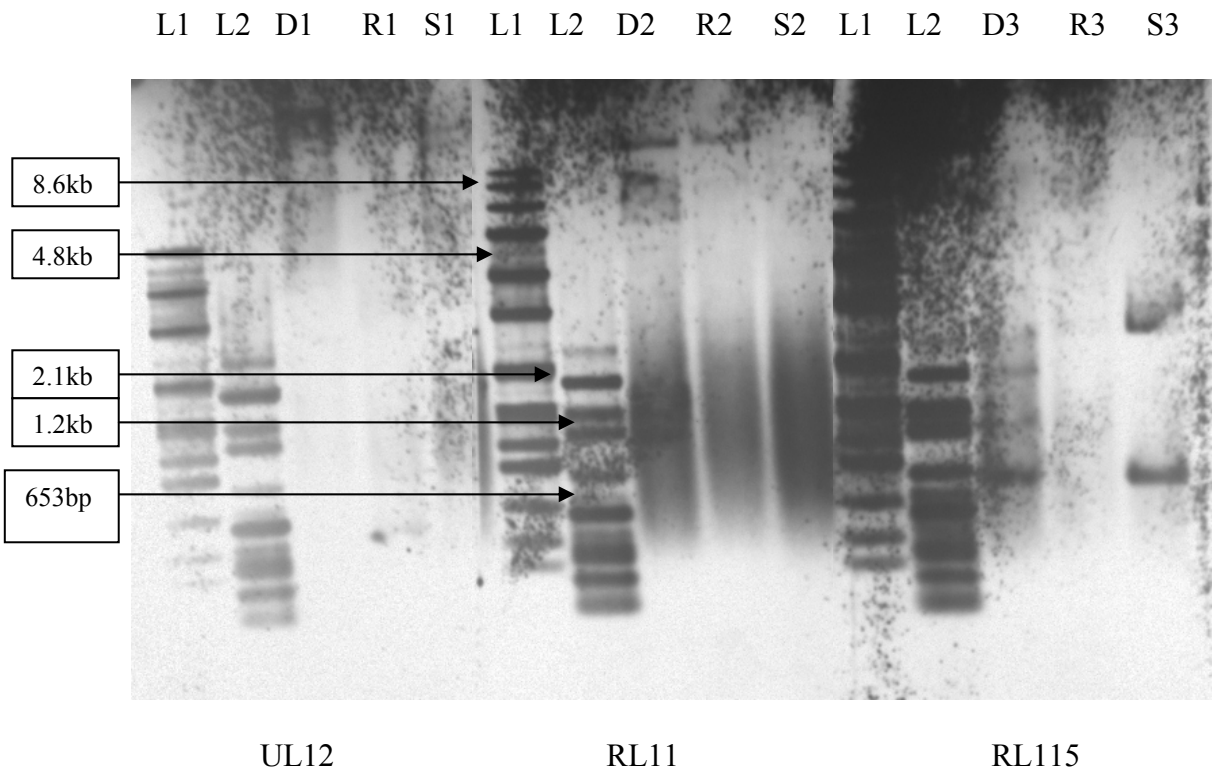


Figure A3. S. blot #4. L1: DIG-labeled molecule weight marker VII (Roche); L2: DIG-labeled molecule weight marker VI (Roche); D: double digest; R:methylation resistant enzyme; S: methylation sensitive enzyme. One microgram of MRC-5 cytoplasmic DNA was electrophoresed after digestion with the following enzymes: BamHI+EagI (D1); BamHI (R1); EagI (S1); BamHI+MluI(D2); BamHI(R2); MluI (S2); BamHI+AgeI (D3); BamHI (R3); AgeI (S3). After electophoresis, the gel was blotted and the filter was hybridized with probes specific for each gene region, (Table 4). Exposure time was 30 minutes.

Table A3. Results for methylation in selected HCMV gene regions based on Southern blot experiment #4.

Gene Region harboring CpG island	Clones overlapping this CpG island in AS orientation	Predicted sizes of fragments in double digest if methylation sensitive sites are methylated	Predicted sizes of fragments in double digest if methylation sensitive sites are not methylated	*Estimated size of fragments generated by from Double digest	Evidence for possible methylation
UL12	pL1016	1877bp 1389bp	488bp 228bp 167bp	14kb	-
RL11	pL811	10+kb 4898bp 1310bp	1145bp	10+kb	+
UL115	pL449	2110bp 1708bp	867bp 841 bp	1.7, 1.1 kb; 653bp	-